MayaPS: Maya hieroglyphics with (LA)TeX

Bruno Delprat and Stepan Orevkov

Abstract

We present a system for hieroglyphical composition of ancient Maya texts, to be used for their palaeography, the production of dictionaries, epigraphical articles and textbooks. It is designed on the base of TeX and PostScript using the Dvips interface, and includes a set of Maya fonts.

The ancient Mayan writing system is very particular: base writing signs attach to each other from all four sides (left, right, top, bottom), and are also rotated and rescaled. This cannot be produced with TeX's usual tools.
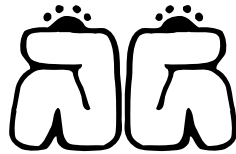
For example, we can type:

`\maya{li.AM2 u.TUN/CHU uj.UJ.ki death.KIMI/la}`

to obtain

 (Dresden codex).

1 Introduction

The present package MayaPS is designed for editing the palaeography of ancient Maya hieroglyphical texts using TeX or LaTeX *and* Dvips. The PhD dissertation [1] and a previously published Spanish language symposium communication [2] are typeset using it. MayaPS is available from `http://picard.ups-tlse.fr/~orevkov`.

To get the above Maya word *xib* (male), we typed:

`\mayaSize{2cm}\maya{422.422}`

As another example, to get *katun* (calendar cycle of 20 years), we input:

`\maya{(023.153.023):220} \emph{katun} (calen...`

2 Structure of the ancient Maya script

2.1 General principles

The ancient Maya logo-syllabic writing has been in use in Central America's Southern Mexico, Guatemala, Belize, Honduras and Costa Rica for more than 1300 years, from the 3rd century AD to the mid-16th century, when Spaniards forbade its use and burned Maya books on religious and political grounds.

This impressive civilization left rich inscriptions on monuments, ceramics and divinatory almanacs. They constitute nevertheless a small volume of available texts: three surviving manuscripts (the Dresden, Madrid and Paris codices) and about a thousand short inscriptions. Maya texts are now largely deciphered, with a variable degree of reliability.

The writing system comprises more than 500 base signs called glyphs. Since the end of the 19th century, Western scholars have set up catalogues of Maya hieroglyphics with different encoding numbers, the most popular being the Thompson Catalogue [11].

Figure 1: Dresden codex page 30b(2) & text palaeography with translation below

400/010.030	+176/204.031	117.260	133/111.023
tsel-ah	*lakin*	*chac-xib*	*kabil*
Was standing	East	red man	sweet

|			
423/515	530.112	515/504.013	026.401
cehel-uah;	*Chac*	*hanal*	*u-bool*
deer tamal	god Chac	meal	its tribute

Ancient Maya words are composed by attaching together primitive (non-decomposable) glyphs, in a way rather similar to how Chinese characters are composed. Composed primitive glyphs are rescaled so that they harmoniously fill a rectangle of a fixed size (called a *cartouche* by Mayanists). The cartouches are placed in a regular way on a page. Maya manuscript texts are organized in blocks of 2 to 16 cartouches which constitute as many sentences, often followed by associated numbers, dates and eventually a picture. According to the number of cartouche spaces available on the almanac page to write a short or long sentence, the scribe would squeeze in or spread out writing signs among the cartouches to avoid empty boxes and obtain a nice-looking page layout.

2.2　Glyph types and orientations

In the ancient Maya writing system, there are two types of primitive glyphs called *central elements* and *affixes*. Usually the shape of central elements is closer to square whereas affixes are narrower.

Central elements always appear in the same orientation but affixes turn so that they stick to other glyphs by their long side, following a general orientation rule.

Complete glyphic cartouches are made up of one to five basic signs or glyphs. Thompson [11] has shown that affixes, like 　 *ni*, present rotation patterns and symmetries, around a central element, such as 　 *KIN*, whose orientation is fixed.

T116 　 *ni*　　T87 　 *te*
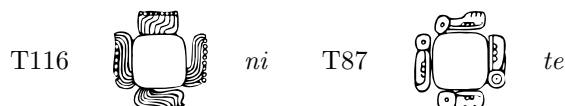
Through analysis of the Maya codices, we have determined that affix patterns follow a definite rule. For example, 　 *te* (tree) is an affix and it usually attaches to a central element like this: 　. So, there are five standard orientations for each affix: when it is single and when it attaches from the left, from the right, etc.

In the Dresden Codex we find the following corresponding cartouche compositions:

Affixes are generally syllabic value signs which can combine together or with a central element to write a Maya word. Central elements are generally

of a logographic nature, corresponding to a morpheme or a word, and read globally; for example 　 *KIN* (sun, day).

A complete glyphic cartouche often corresponds to a lexical entry with preceding and following grammatical affixes as in 　 *KIN-ni* (sun, day), but it can also in some cases correspond to two words if they are short, or more rarely otherwise to a part of an expression spelled over two cartouches.

## 2.3　Glyph composition into cartouches

For the composition of glyphs the following standard notation is used in the historical and linguistic literature on ancient Maya: if 　A　 and 　B　 are two glyphs (primitive or not), then `A.B` and `A:B` encode the glyphs 　A B 　 and 　$\frac{A}{B}$　; for example: 204.031 　 and 204:031 　. To control the order of composition, one can use parentheses in the same way as in mathematical formulas. For example, both `A.B:C` and `A.(B:C)` stand for 　$\begin{array}{c}A \; \frac{B}{C}\end{array}$　 but `(A.B):C` stands for 　$\frac{A \; B}{C}$　; thus: 026.172/023 　 and (154.123)/306 　 (in glyph codes, '`/`' means the same as '`:`').

3　Description of MayaPS

3.1　Main features

The text cartouches 　 and 　 that we used above are composed of these primitive glyphs:

MayaPS does not consider any grammatical function or linguistic meaning of primitive glyphs. They are just graphical elements which are the elementary bricks of Maya typesetting, like letters in European languages.

As should be clear already, each primitive glyph is referred to by its code (more specifically called the *glyph code*). The glyph codes for the above are `422`, `023`, `153`, and `220`. In general, a glyph code is any sequence of digits `0...9` and letters `a...z`, `A...Z`. The encoding system is rather flexible. For example, after the command `\mayaDefine{A9z}{422}` you can type `maya{A9z}` to get 　.

Any formula of this kind is allowed by MayaPS, even something like this:
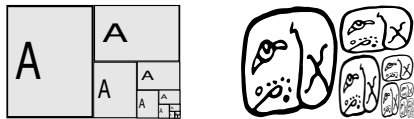
The picture on the right hand side (a fantasy writing cartouche, as there is no such glyph in ancient Maya writing) is printed by the command

```
\maya{322.322:(322.322:(322.322:(322.322:(322.
322:(322.322))))}
```

The type of each glyph (affix or central) and the five default orientations for each affix are written in the font file. Orientation can be changed with *modifiers* `-|+*?` whose meaning is:

The modifiers can be composed together yielding:

Let us discuss again the glyph . We see that the primitive glyph `023` occurs here in two different ways: and . Moreover, in it occurs twice horizontally. MayaPS automatically chooses the orientation of ea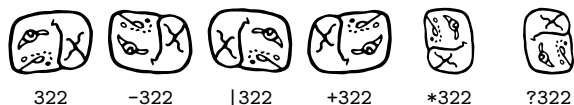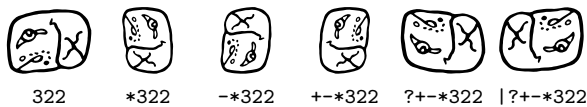ch primitive glyph of the affix type according to "orientation rules" formulated by the first author after a careful analysis of ancient manuscripts. Of course, these rules have exceptions, but it is easy to handle them. For example, if you type `\maya{422.222/024}`, you obtain (the default orientation), but if you type

`\maya{-422.410}`, you obtain .

A more representative example is the palaeography of page 22c of the Dresden Codex, due to the first author:
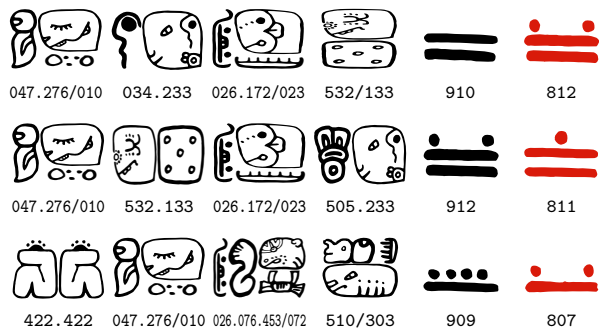
To obtain this, we typed:

```
\mayaC{  % \mayaC = glyphs with input codes
047.276/010 034.233 026.172/023 532/133 910 812
047.276/010 532.133 026.172/023 505.233 912 811
422.422 047.276/010 026.076.453/072 510/303 909
                                              807}
```

MayaPS permits the support of multiple fonts. In this paper we use mostly the font `codex` created using the tools mentioned in §4.1, but another style glyph set `gates` has been implemented based on the same glyph codes, for example:
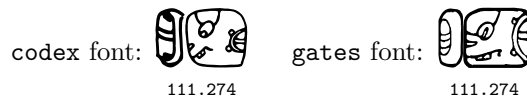
MayaPS provides a tool to add or replace glyphs in existing fonts and a tool for making new fonts.

3.2 Substitutions (ligatures)

A list of substitutions is associated with each Maya font. As we mentioned in the introduction, a new substitution $s_1 \rightarrow s_2$ can be defined by the command `\mayaDefine{`s_1`}{`s_2`}` where s_1 and s_2 are any strings (chains of characters). Substitutions are applied to the arguments of glyph drawing macros. They are applied non-recursively. Some substitutions can be predefined in a Maya font. Three types of substitutions are predefined in the font 'codex':

(1) *Ligatures.* One or several affixes or central elements can be melded inside a central element or, more frequently, inside a head figurative element instead of being simply attached to it, forming a *ligature* as a single bound form. For example, when you

type `\maya{070:349}`, you obtain `353` rather than `070:349` because of the predefined substitution (ligature) `070:349` → `353`. Here we typed `\maya{070:(349)}` to print the non-ligatured form. The ligature was not applied because '`070:349`' is not a substring of '`070:(349)`'.

As another frequent ligature example you have: `373` *Cacau* D7c (2), that decomposes into simpler glyphs: `369<023/023>`. The operator `<...>` indicates that both affixes `023` are placed in the centre of `369` .

Within Maya texts, both forms — melded into a ligature (single glyph code), and separately drawn (2–3 glyph codes) — are equivalent and may constitute orthographical variants, e.g.: and . Our catalogue includes around 100 ligature glyphs.

(2) *Thompson codes.* The basic glyph codes in the font 'codex' are based on glyph numbers of

the Evreinov catalogue [3], from which font drawings were derived. However, many specialists are more familiar with glyph codes in Thompson's catalogue [11]. Due to predefined substitutions T1 → 026 , T2 → 410 etc., those codes can be used also for text input.

(3) *Phonetic values.* Reconstituted phonetic values in the Maya language can also be used to ease text input. Phonetic values of affixes are conventionally written in lowercase letters: a → 050 , aj → 044 , ak → 506 etc.; for central elements uppercase letters are used: AT → 200 , BA → 213 , BA2 → 212

As with Thompson codes, predefined substitution tables permit the use of multiple character input methods, analogous to Chinese character computer input with either *pinyin* (PRC's official romanization), *cangjie* (decomposition into graphic keys) or *dianbaoma* (Chinese telegraph codes).

4 Maya fonts

4.1 Font creation mechanism

A MayaPS font is an ASCII text file with the extension `.mpf`. Its structure is rather flexible. It is described in detail in [8]. It has several sections of PostScript code (a header and glyph definitions) [6] separated by lines starting with `%@`. TeX macros use these marks to select needed sections for including them into '`mayaps.tmp`' (see §7.1). Substitution rules (see §4.2) have the form `%L@` s_1 s_2.

There is a tool (involving a special vectorizer '`cotrace`') for creating MayaPS fonts out of monochrome bitmaps. The fonts supplied with MayaPS are made with it.

When MayaPS fonts created with this tool are used, they generate Type 1 fonts [5] in the resulting `ps` file. As Type 1 Maya fonts are used, the resulting `pdf` document after conversion is considerably smaller than the intermediate `ps` file. Only definitions for those font signs used in the text are included, and just once.

The `\mayaAddGlyph` macro allows for inclusion of a new glyph from an `eps` file but we do not recommend using it too frequently because it rapidly increases the resulting `ps` and, especially, `pdf` files.

4.2 Currently available fonts

To date, three extensive Maya script fonts and a partial Olmec script font have been produced by the first author.

The font '`codex`*'* was designed primarily from drawings of the Evreinov glyph catalogue [3] and is the one used up to now in this article.

The font '`gates`*'* is derived from the lead cast font designed by William Gates in the 1930s for his book [4], and has been implemented based on the same glyph codes as for font '`codex`'. For example, the beginning of the above quote from the Dresden Codex printed in the font '`gates`' looks like:

451.452 026.314/(314) 111.274 047.276/010 913 810
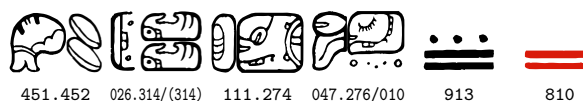
(we typed `\gates` and then, just copied the same codes used above for a text in the '`codex`' style).

The font '`thompson`*'* was made based on the Thompson catalogue [11] drawings, where the original Txxx codes from the catalogue are implemented, showing that glyph codes in different fonts can be independent.

A correspondence ligature table, included in the font, supports glyph input using Evreinov-derived codes in a text displayed in the '`thompson`' font, as shown:

T24.T1047a T1.T19:T59a T58a.T1026.T103 T15.T736a:T140 XVIn

Similar correspondence ligature tables are included in the '`codex`' and '`gates`' fonts for glyph input using Thompson codes and phonetic reconstituted values, for example:

T24.T1047a T1.T19:T59 T58.T1026.T103 T15.T736a:T140 XVIn
111.274 026.144:056 505.233.112 047.276:010 916

li.AM2 u.TUN/CHU uj.UJ.ki death.KIMI/la

The font '`olmeca`*'* is derived directly from drawings of *La Mojarra* stella [7] of the 2nd century AD and from the *Tuxtla* tablet. It represents a partial set of Olmec glyphs, which were composed in vertical texts without rotation of affixes.

Bruno Delprat and Stepan Orevkov

5 Principles we tried to follow

5.1 Length optimization

Suppose you already have a `ps` file (produced by TeX/Dvips [9]) in an alphabetic language and you include ancient Maya glyphs into it. Then MayaPS adds to the `ps` output only:

- MayaPS header (7 Kb);
- definitions of primitive glyphs (0.5–3 Kb per glyph for 'codex');
- about 60 bytes for each occurrence of each composed glyph.

MayaPS includes the definitions of only those primitive glyphs which are effectively used in the text. Each definition is included only once even if the primitive glyph is used many times. This property holds after conversion from `ps` to `pdf`, because Type 1 fonts are used for primitive glyphs (as usual, the size of `pdf` generally lies between the sizes of `ps.gz` and `ps`).

5.2 Simplicity of installation and no need of support

To use MayaPS, it is enough to copy a few files into any directory (folder) 'visible' by TeX, for example, the directory where the `tex` file is. In particular, no extra font in the usual sense is needed (a typical beginner's problem is how to make TeX 'see' a new font).

The algorithm to draw a composed glyph is implemented in the PostScript language [6], and features of Dvips [9] are used for calling it from a TeX file via a `\special` macro (this is why MayaPS does not work with pdfTeX). So, TeX, Dvips, and PostScript are needed. Nothing else is used in MayaPS.

The only exception is the tool for creating new MayaPS fonts (`mpf` files) where C programs are used, but the font file format is described and it is easy to make an `mpf` file out of a Type 1 font (detailed instructions are given in reference [8]).

6 A few words about the implementation

6.1 Interaction between TeX and PostScript

A glyph code (example: 111.+176/111 for) is passed to the `ps` output by the Dvips command:

    \special{"M(111.+176/111) w h d E}

where $w \times h$ is the cartouche size and d is the font descriptor (an integer number). Dvips literally includes the argument of `\special{" }` into the `ps` file and the task of drawing the composed glyph is delegated to a PostScript interpreter. The glyph

drawing subroutine `E` is defined in the header included to the `ps` file by the Dvips command:

    \special\{header:mayaps.tmp}

(see [9]; §5 for more detail).

Before issuing the command `\special{"M...E}`, all primitive glyph names are extracted from the glyph code and checked to see if their definitions are already included into the header `mayaps.tmp`. The token list `\output` is extended so that at the end of each page the definitions of all newly-appeared glyphs are copied from `mpf` files to `mayaps.tmp`.

6.2 Substitution mechanism

In earlier versions of `MayaPS`, the substitution mechanism was implemented by creating for each substitution $s_1 \rightarrow s_2$ a macro whose name (control sequence) contains s_1 and whose expansion is s_2.

Then, for each substring of each composed glyph, the corresponding macro was checked for existence by this command:

    \ifx\csname ... \endcsname\relax

However, this command leaves the tested control sequence in TeX's memory forever. As a result, TeX's usual capacity (60000 control sequences) was exceeded when the thesis [1] exceeded 300 pages.

The new substitution mechanism creates the tree of initial subwords of left hand sides of all substitutions. Now the number of control sequences used does not exceed the size of this tree.

Another attempt to adapt (LA)TeX for ancient Native-American languages that concerned Olmec writing was done in [10], using a very different approach from ours.

References

[1] B. Delprat. *Le codex de Dresde: Paléographie et traduction comparée d'un almanach maya du 15e siècle.* Thèse de doctorat, Institut National des Langues et Civilisations Orientales, Paris, n.p. (thèse en cours).

[2] B. Delprat and S. Orevkov. mayaTeX – un sistema de composición tipográfica de textos

jeroglíficos mayas para la computadora. In
*XXI Simposio de investigaciones arqueológicas
en Guatemala*, Guatemala de la Asunción,
23–27 July 2007.

[3] E. V. Evreinov, Yu. G. Kosarev, and
V. A. Ustinov. *Primenenie elektronikh
vychislitel′nykh mashin v issledovanii
pis′mennosti drevhikh maiya [The Use of
Electronic Computing Machines Applied
to Research on Ancient Maya Writing]*.
Akademia nauk SSSR [Academy of Sciences of
the USSR], Novosibirsk, 1969. 4 vols.

[4] William E. Gates. The Dresden Codex
reproduced from the tracings of the
original colorings and finished by hand. In
Maya Society Publication, number 2. The
Maya Society at the Johns Hopkins University,
Baltimore, 1932.

[5] Adobe Systems Inc. *Adobe Type 1 Font
Format*. File `T1Format.pdf` available on
`http://www.adobe.com`.

[6] Adobe Systems Inc. *PostScript Language
Reference Manual*. Files `plrm.pdf` and
`plrm2.pdf` available on `http://www.adobe.
com`.

[7] Martha J. Macri and Laura M. Stark. *A
Sign Catalog of the La Mojarra Script*.
Pre-Columbian Art Research Institute, San
Francisco, 1993.

[8] Stepan Orevkov. *MayaPS: Typing Maya with
TEX/LATEX. Reference manual.* available on
`http://picard.ups-tlse.fr/~orevkov`.

[9] T. Rokicki. *Dvips: A DVI-to-PostScript
Translator*. File `dvips.pdf` included
in most TEX distributions, available on
`http://www.ctan.org`.

[10] A. Syropoulos. Typesetting Native American
languages. *Journal of Electronic Publishing*,
8(1), 2002.
`http://www.press.umich.edu/jep`.

[11] J. E. S. Thompson. *A Catalogue of Maya
Hieroglyphs*. Univ. Oklahoma Press, 1962.

⋄ Bruno Delprat
   INALCO & SeDyL-CNRS, 7, rue
      Guy Môquet
   94801 Villejuif cedex
   France
   `brunodelprat (at) club-internet
      dot fr`
   `http://celia.cnrs.fr/Fr/Labo/
      Delprat.htm`

⋄ Stepan Orevkov
   Institut de mathématiques de
      Toulouse, Université Paul
      Sabatier
   31062 Toulouse
   France
   `orevkov (at) math dot ups-tlse
      dot fr`
   `http://www.math.univ-toulouse.
      fr/~orevkov/mayaps.html`